

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 21:32:49 : Search time 3319.91 Seconds
(without alignments)
11301.897 Million cell updates/sec

Title: US-09-719-017a-2

Perfect score: 1793

Sequence: 1 gattcccccgttgacacatca.....caattactcaatgcgcgcg 1793

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_others:*

33: em_htgc_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	1793	100.0	1793	6	AX008753	AX008753 Sequence
2	1793	100.0	1793	6	AX025996	AX025996 Sequence
3	1132.8	63.2	1727	1	AFMAANL	D13419 A. faecalis
4	1097.6	61.2	1727	6	E07629	E07629 gDNA encodi
5	500	27.9	7508	1	ECORGNB	J01695 E.coli rRNA
6	500	27.9	7508	1	ECORGNB	V00348 E. coli rib
7	500	27.9	11679	1	AE000471	AE000471 Escherich
8	500	27.9	11918	12	CUV72488	U72488 Cloning vec
9	500	27.9	176195	1	ECOM89	U00006 E. coli chr
10	421	23.5	860	12	SYNRGNABP	K00764 e.coli rna
11	404	22.5	5094	12	AX013859	U13859 pRK232-8-cl
12	404	22.5	6733	12	AF216802	AF216802 Shuttle v
13	404	22.5	6984	12	AF216803	AF216803 Shuttle v
14	404	22.5	8565	12	SYNPR610T	M29896 Cloning vec
15	404	22.5	9113	12	CYPCG1408	U21228 Promoter-pr
16	404	22.5	13331	12	AF209190	AF209190 Shuttle v
17	403.6	22.5	502	23	E09023	E09023 3'-terminal
18	401.6	22.4	7965	12	AF327719	AF327719 Cloning v
19	401.6	22.4	11299	12	AF327720	AF327720 Cloning v
20	396.4	22.1	2491	1	ECOBTRA	M10123 E.coli b1ra
21	388.4	21.7	4204	6	E02679	E02679 All sequenc
22	388.4	21.7	4207	6	E02104	E02104 DNA sequenc
23	388.4	21.7	4222	6	E02098	E02098 DNA sequenc
24	388.4	21.7	4466	6	E02001	E02001 DNA sequenc
25	388.4	21.7	4466	6	E06722	E06722 Plasmid pTP
26	369	20.6	7416	6	AX354080	AX354080 Sequence
27	336	18.7	5641	6	AX113748	AX113748 Sequence
28	336	18.7	5670	6	AX113749	AX113749 Sequence
29	336	18.7	5826	6	AX113746	AX113746 Sequence
30	336	18.7	6071	6	AX113747	AX113747 Sequence
31	332.8	18.6	4176	6	AF028838	AF028838 Cloning v
32	332.8	18.6	4176	6	AX13038	AX13038 Artificial
33	332.8	18.6	4176	12	SYNPRC99A	M22744 Cloning vec
34	332.8	18.6	4176	12	AX013872	U13872 pTic99a clo
35	332.8	18.6	4214	6	AX027488	AX027488 Sequence
36	332.8	18.6	4239	6	A29289	A29289 pSEC-Bpl co
37	332.8	18.6	4357	6	AX072889	AX072889 Sequence
38	332.8	18.6	4476	6	AR102989	AR102989 Sequence
39	332.8	18.6	4476	6	AR172783	AR172783 Sequence
40	332.8	18.6	4566	12	AF177933	AF177933 Cloning v
41	332.8	18.6	4593	6	AR102988	AR102988 Sequence
42	332.8	18.6	4593	6	AR172782	AR172782 Sequence
43	332.8	18.6	4593	12	CVPRK2332	X70478 Plasmid PKR
44	332.8	18.6	4593	12	XXU02439	U02439 Cloning vec
45	332.8	18.6	4635	12	AF177932	AF177932 Cloning v

ALIGNMENTS

RESULT 1

AX008753

LOCUS AX008753

DEFINITION Sequence 2 from Patent WO9964607.

ACCESSION AX008753

VERSION AX008753.1 GI:9996244

KEYWORDS

SOURCE

ORGANISM

synthetic construct.

synthetic construct

artificial sequence.

1 (bases 1 to 1793)

REFERENCE

FAVRE-BULLE O., GUILTON C. and PIERRARD J.

AUTHORS

JOURNAL

FAVRE-BULLE OLIVIER (FR); GUILTON CAROLE (FR); PIERRARD JEROME (FR); RHONE-POULENC NUTRITION ANIMAL (FR)

TITLE

Industrial method for producing heterologous proteins in e.coli and strains useful for said method

FEATURES

source

1..1793

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="cassette d'expression"

123..1193

CDS

```
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC07435.1"
/db_xref="GI:9969245"
/translation="MQTRKIVRAAAVQAASPNYDLANGVDKTIELARQARDEGDLIV
FGEWMLPGYPRPHVWLGAPWMSLKSARYANSLSIDSSEFRIQAAATGILFIALGY
SERGGSLYLQOCLIDDGOMLSMRKLKPHVEVTEVGEVARDLIYSDPELRVGA
LCQNEHLSPLSKVALYSOHEAIHTAAMSFSLYSOAHALSAKUMMAISOIYSEVGC
FTTAASVYQDETLDMLEVBHNSILKVGSSMIFAPDQRTIAPYLPHPAEGLITA
DLNMEIIFAFAINDPVGHSKPEATRIYLDLGRHREPTRVHSSVIDEAPRPHVDS
TAAPVAAVSQTDSDTLVQEPS".
BASE COUNT      412 a      527 c      478 g      376 t
ORIGIN
Query Match      100.0%: Score 1793: DB 6: Length 1793:
Best Local Similarity 100.0%: Pred. No. 0:
Matches 1793: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
```

```
|||||
Db      781 CCATGCGCGCCAGCAGTGTGTCACCCAGSAGACACTGGACATGCTGGAAAGTAGTGAAC 840
Qy      841 acaagcgtccctgtctgtaagaatggtgagcggtgagttcatgtatctttgcccggacggac 900
Db      841 ACAAGCGCTCCCTGCTGTGAAGAGGGCGCGGAGTGCATGTATTTTCCCGCGGAGAC 900
Qy      901 gacactgtctccctcactgcccacacgactgcccgaagcgctgtactctgcgaatccga 960
Db      901 GCACATTGTGCTCTCCACTGCGCACACGATGCGGAAGGCGCTGATCTATTCGCATCTGMAA 960
Qy      961 tggaaagaatgtgctctgcgccaagcgatcaacgacctgtgtggccactactccaaacccg 1020
Db      961 TGGAGAAATTGCTTTCGCGCAAGGCGATCAACGACCTCTGTGGCCACTTCACAAACCCG 1020
Qy      1021 agggccaccgtctgttactggaacctggggggaacctgtgaagccatgactgggtgtactca 1080
Db      1021 AGGCCACCGCTGTGTACTGTGACCTGGGGCACCGTAGAGCCATGTGACTGGGGTCAATTCCA 1080
Qy      1081 aaagcgtgataccagaagaagctcccgagccgacgtgcaaaagtaacggtcgtgcgcgtcg 1140
Db      1081 AAAGCGTATCCAGSAGAAAGCTCCGAGCCGCGACGTCGAAMAATGAGGCTGGCGCGTGG 1140
Qy      1141 ccgtcaagccagactcaaggaactcgtgaatcgtgtgtcaagaacggtccgtgaacccaaa 1200
Db      1141 CCGTCAGCCAGACTCAGAGACTCGGATACGCTACTGTGCACAAACCGTCTGACCCCAA 1200
Qy      1201 agatgaagaagccgggggaactgtccgggtcttgatctctctgcgcccggatccact 1260
Db      1201 AGATGACAAAGGCCCGGCANAACTGTCCGGTCTTATTTCTCTCTGTCGTCGGATCCACT 1260
Qy      1261 agttctagaagtcgaactcgaagcaltgcaagcttgggtcccaactgaccccaatgcgaact 1320
Db      1261 AGTTCTAGAGTGTGACCTGACGACTGACAAAGCTTGGGTCCACCTGACCCCAATGCCAAG 1320
Qy      1321 cagaagtgaaacgcgtgtgagcgccgactgtgtagtgtgtgtgtcttcccatgctgaagta 1380
Db      1321 CAGAAGTGAACGCGGTGCGCGCATGTGATGTGGGGTCTCCCATCGACGAGTAGGA 1380
Qy      1381 actgcgaagcatcaaatgaagaagcctcagtcgaagaagcttgggtctctgtcttactc 1440
Db      1381 ACTGCGACGACATCAATTAACGAAGAGCTCAGTGCAGAAAGACTGGCGCTTGTTTATTC 1440
Qy      1441 tctgtcttgctgggtgaagcgtctcctgaagtaagacaatccgcggagcggaatttgaaac 1500
Db      1441 TCTTGTGTTGTCGCTGTAACGCTCTCTCTGAGTAGACAAATCCCGCGGAGCGGATTTGAAC 1500
Qy      1501 gttgcgaagaacacggcccggaaggtgtggtgggaggaagcccgacataaactgccagcat 1560
Db      1501 GTTGCAGAACACGGCCCGGAGGGTGGCGGAGACGCGCCGCAATTAACCTGCCAGGCAT 1560
Qy      1561 caaatlaagcagaagacalccctgaacgatagtgctcttctgtcttctcaaaactctcct 1620
Db      1561 CAATTAAGCAAGAGGCCATCTCTGACGATGGCTTTTGGCTTTCATCAAACTCTTCCT 1620
Qy      1621 gtcgtcatactacaagccatcccccacagataacggttaactagctcgtctttgtgacac 1680
Db      1621 GTCTGTATATTCTACAGAGCATCCCGCCACAGATAGGTAACCTGCTGTTTGTGATC 1680
Qy      1681 aggaagaagcagctatgaacacactccttaaaacccctggaacacacttgggaatgataat 1740
Db      1681 AGGAAGAAGCAGCTATGAAACCACTCTTAHAACCTTGGAACACTTTGGCATTTGATATANT 1740
Qy      1741 gctcagacatgtatgtgcccgaagaagacaaacttaactaatgaccgcg 1793
Db      1741 GCTCAGCAATGTGATGTGTCGGAAGACGAACAACATTAATCAATGCCCGCG 1793

RESULT 2
AX025996      1793 bp      DNA      linear      PAT 16-SEP-2000
LOCUS      DEFINITION Sequence 1 from Patent FR2787121.
ACCESSION AX025996
```

VERSION	AX025996.1	GI:10187454
KEYWORDS	Alcaligenes faecalis.	
SOURCE	Alcaligenes faecalis	
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;	
REFERENCE	1 (bases 1 to 1793)	
AUTHORS	Favre,B.O., Plerriard,J. and Battisse,D.N.	
JOURNAL	Patent: FR 2787121-A 16-JUN-2000;	
FEATURES	AVENTIS CROPS/SCIENCE (FR)	
source	Location/Qualifiers	
	1..1793	
	/organism="Alcaligenes faecalis"	
	/db_xref="taxon:511"	
	123..1193	
CDS	/note="unnamed protein product"	
	/codon_start=1	
	/transl_table=11	
	/protein_18="CAC09069.1"	
	/db_xref="GI:10187455"	
	/translation="MQTRKIVRAAIVOAAENPVDATGVKDTIELANQARDEGDLIV FGEHMLPGYPRPHWILGAPAMSLKRSARYANSLSDSAEFORIAQARTGIFALGV SERGSGSLYLIOCLIDKGGOMLRRLKIPHYRYTVEGVDLIVSDPELRGVA LCQEHSLPSLKVALYISQHEAIIHTAARPSLTYSQAHLSAKNNMAISOIYSEGGC FTTAASSVYDETIDMLEVEGHNSLKLKVGGSMMTRAPDRITAPLUPHAEGLIIA DLNDEETFAKAIINDPVGHYSKPEATRIVLVDLGHREPTRVHSKSVIOEAPHPVOS TAAVVAVSOTODSDPTLIVQEPS"	
BASE COUNT	412 a	527 c 478 g 376 t
ORIGIN		
Query Match	100.0%;	Score 1793; DB 6; Length 1793;
Best Local Similarity	100.0%;	Pred. No. 0;
Matches 1793; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY	1	gaatccctgttgaacaataatcatcgcgaactagttaaactagtcgcagcttgctgcagg 60
Db	1	GAATTCCCTGTTGAAATAATTAATCATCGAAGTAAGTAACAGACACCTGGCTGCAGG 60
OY	61	tgcacctgcagccaagcttggcgctacacatccaatattgtatctcctaaggaatacttac 120
Db	61	TGCACCTGCAGCCAAGCTTGGCGTACATTCATTAATTGTATTAAAGAAATACCTTAC 120
OY	121	atactgcagacaagaanaaaatcgtctcggcggaagccgcctgatacaggcgcgcctctcccaactacg 180
Db	121	ATATGCAGACAAGAANAATCGTCCGGGACGCCGCTGACAGGCCGCTCTCCACACTACG 180
OY	181	atctggtcaaacgggtgttgaataaaccaattgagctgctgcgttcagagcccgatgaggct 240
Db	181	ATCTGGCAACGGGTTTATATAAAMCAATTGACCTGGCTGCTAGGCCGCCGCTGAGGGCT 240
OY	241	gtgaactatcgtgttggtagaacctcgctgcgccggtatcccttcgaagctcgtgcctcg 300
Db	241	GTGACTATATCGTGTGGTGTGAACCTGCTGCCGCTATCCCTTCCACAGCTTGGCTGG 300
OY	301	gcgcaacggcgctgtgcgtgtgaataacaglygcgcgtactatgccaactcgcgtctcgtcg 360
Db	301	GGCGACCGGCGTGTGGTGAATAATGCAATGCGCCGCTACTACTGCAACTGCGCTCGTGG 360
OY	361	acagtcgaggaftttaaagcatatgcccaagcgccgaagcgaacttgggtatttcaatcgac 420
Db	361	ACAGTCGAGAGTTCATACGCAATGGCCAGGCGCGACGACACTTGGGTATTTTTCATCGCAC 420
OY	421	tgaggtataagcgaagcgcgagcgcaagccttacctggccaatgaccttgatctcaagaca 480
Db	421	TGAGGTATAGCGAGCGGACGCGGCGGACACTTTTACTTGGGCCAATGCTGATTCGACGACA 480
OY	481	agggcacagatcgtgtgtcgcgctgcgcaaaactacaactacatactgttttagcgcaacgctgt 540
Db	481	AGGCGCAAGATCCTGTGTGCGCGTCCCAAACTCAACATGTTGAGCGCACCGCTGT 540
OY	541	ttgtggaaggttatgcccgaagatctgattgtgtccgaacaacggaagcttgccggtcggtg 600

Db	541	TTGTGGAAAGGTTATGCCCCGAGATCTGATTTTGTGCCACACCGAGCTGGGCGCGTGGTG	600
Qy	601	ccctgtgtctgtggaagcactgttccccccttgagcaagtacgcgtgtactcccaagc	660
Db	601	CCCTGTGCTGCTGGGAGACGACTGTGCCCTTTGAGCAAGTACGGCGTGTACTCCAGCAG	660
Qy	661	aagcattcaccattgcccgtgtgcgttcccttctgcgttaagcaggaacagccatgcgc	720
Db	661	AAGCATTTCACATTTGCGCGCGCTGGGCGGCTCTTTTTCGCTGTACAGCGAAGCGCCCATYGGC	720
Qy	721	tcaagcgcaaggtgaacatgtctgtccctcgaaatctatctcgtttgaagccagtttctta	780
Db	721	TCAGCGCCAAAGGTAAACAATGGCTCCTCGCAAAATCTATTTCGGTTGAAGGCCAGTCTTTTA	780
Qy	781	ccatctgcgcgcgaagcagtgctgtcaaccgaagagacacttggacatgtctgtgaagttagt	840
Db	781	CCATGCGCGGCAAGCATGTGTGTACCCAGGAGACACTGGAGCATGTGTGAAGTAAAGTTGAA	840
Qy	841	acaaagctctcccctgtcgtgaagttggcgcgagtcagttccatgtatlttgcgcgaagcag	900
Db	841	ACAAAGCCTCCCTCTGTAAGATGGGCGCGCGAGTTTCATGATTTTGGCGCGAGCGAGC	900
Qy	901	gcaacttggcctccctactcttcaacaagatgcgaagagccttgatctatgtcgaatctgaac	960
Db	901	GCACATTGTGGTCCCTTACTCTGCCACAGATGCGCAAGGCGCTGATCTTGGCATTTGCAACA	960
Qy	961	tggaaagaattgtccttcgcgaagcgatcaaacagccctgtggccactactccaacccg	1020
Db	961	TGGAAGAAATTTGCTTTCGCCAAGGCGATCAAGCAGCCTGTGGGCGCACTACTCCAACCG	1020
Qy	1021	aggccaccgctctgtgtacttggaccttgggacccgtgaagcccatgtactcgggtgtacca	1080
Db	1021	AGGCCACCCGCTGTGTACTGTGACTGTGGGCGACCGTGAAGCCATGTACTCGGTGTCTCA	1080
Qy	1081	aaaggtgtatcccaagaagaagctcccgagccgaagctgtcaaaagtacgcttgcgcgcgcg	1140
Db	1081	AAAGGTGATCTCAGGAAGAAGCTCCGAGCGGACGTCGCAAGATAGGTGGCTGGCGCGTGC	1140
Qy	1141	ccgtcaagcacaagcacaagactcggatcacgtactctgtgtcaagaagacgctctgaaccaca	1200
Db	1141	CCGTACGCGACACTCAGAGACTCGGATATCGTACTGTGTGCAGAGACCGTCTGACCCCAAA	1200
Qy	1201	agatgaacaagcccggaacaaactgtccggttcttgaatctcttctgcgtcccgatccact	1260
Db	1201	AGATGACAAAGCCCGGCGCAAACTGTCCGGGTTTGATTCCTTCTGGGTCCCGATTCAC	1260
Qy	1261	agttctagatgcgcagcctcgcagagcatgtcaagcttgggttcccaactgaaccatgtccgaact	1320
Db	1261	AGTTCTAAGTGCAGCTGCGAGGATGCGACAACTTGGGTGTCCACCTGACCCCATGCGGA	1320
Qy	1321	cagaagtbaaacgcgcgttagcgcgaatgtgtatgttgggtctcccaatgtgaagatagga	1380
Db	1321	CAGAAGTCAAAAGCGCGTATGCGCGCGATGTGTGTGGGTTCTCCCATGTGAGAGATGAGGA	1380
Qy	1381	actgcagagcaacaataaaacgaagaagctcagtcgaagaagcttggccttctgttattc	1440
Db	1381	ACTGCGACAGCATCAAAATAAAACGAAGAGCTCATGCAAAAGACTGGGCGCTTGTGTTAT	1440
Qy	1441	tgttgttgtgtgtgaaacgctctcttcctgagtgaagacaacatccgcgggagcggatltgaac	1500
Db	1441	TGTTGTTGTGTGTGTAACGCTCTCTCTGTAGTAGGACAAATTCGCCGAGACGGATTTGTA	1500
Qy	1501	gttcggaagcaagcccggaagggttggcgggacgaagcgcgcgcgaataacttgcgaagat	1560
Db	1501	GTTGCGAAGCAGCGCCCGAGGGGTGGCGGCGAGCGCCGCCATTAATCTGCCAGGAT	1560
Qy	1561	caaatlaagcagaagccatctctgtacggaatgtggccttltgcgtttctcaaaacttctct	1620
Db	1561	CAAAATTAAGCAGAAAGCCCATCTGTACGAGATGGGCTTTTGGCTTTCTAACAATCTTCT	1620
Qy	1621	gtgtctatatctacaagcatalccccccaagataagcgtbaaactagcctcgttltgtcatc	1680
Db	1621	GTGTCTATATCTTACAACCTATCCCGCCACAGATACGTTAACTATACCTCTGTTTGTGATC	1680

QY 1681 aggaagcagcctatgaaccctccttaaaacccctgagacacattgcatgatacat 1740
|||||
Db 1681 AGGAAGCAGCCTATGAACCACTCTTAAACCCCTGGAAACATTGGCATTTGATCATTAAT 1740
QY 1741 gctcagcagcattatgctgcggaagagcaacaactactcaatgcgcgcg 1793
|||||
Db 1741 GCTCAGCAGCATTGTATGTGCGGAAGAGCAACAATTACTCAATGCCCCGCG 1793

RESULT 3
AFPAANL 1727 bp DNA linear BCT 03-FEB-1999
LOCUS
DEFINITION A. faecalis DNA for arylacetoneitrilase, complete cds.
ACCESSION D13419
VERSION D13419.1 GI:216202
KEYWORDS arylacetoneitrilase.
SOURCE Alcaligenes faecalis (strain:JM3) DNA.
ORGANISM Alcaligenes faecalis
Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
Alcaligenes.
REFERENCE 1 (bases 1 to 1727)
AUTHORS Kobayashi,M.
TITLE Direct Submision
JOURNAL Submitted (16-Oct-1992) Michihiko Kobayashi, Kyoto University,
Department of Agricultural Chemistry, Faculty of Agriculture,
Oiwake-kyo, Kitashirakawa, Sakyo-ku, Kyoto 606, Japan
(Tel:075-753-6114, Fax:075-753-6128)
2 (bases 1 to 1727)
REFERENCE Kobayashi,M., Izui,H., Nagasawa,T. and Yamada,H.
AUTHORS Nitrlase in biosynthesis of the plant hormone indole-3-acetic acid
TITLE from indole-3-acetonitrile: cloning of the Alcaligenes gene and
JOURNAL site-directed mutagenesis of cysteine residues
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 90 (1), 247-251 (1993)
COMMENT 93126352
Submitted (16-Oct-1992) to DDBJ by:
Michihiko Kobayashi
Department of Agricultural Chemistry
Faculty of Agriculture
Kyoto University
Kitashirakawa-oiwakecho
Sakyo-ku
Kyoto 606
Japan
Phone: 075-753-6114
Fax: 075-753-6128.
Location/Qualifiers
FEATURES
source 1..1727
/organism="Alcaligenes faecalis"
/strain="JM3"
/db_xref="taxon:511"
304..1374
/codon_start=1
/product="arylacetoneitrilase"
/protein_id="BA02684.1"
/translation="MOTRKIVRAAAVQAASPNYDLYATGVDTKTELARQARDECDLIY
FGEIWLPGYFHWLGAAPWASLKYSAKYANSISLDSAEFORIAQAARPLIGFIALGY
SRRSGSLTIGGCLTDDKGOMLSRRKLPKPTHERYVREGVYARQDIYSDTELRVGA
LCCEWHSPLSKYALYSQHEAHIHAMPSEFSLXSDAHLSKYNMAAQIYSVEGQC
PTTAASSVVTQETLDMLEVEGHNASLTKVGGSSMTFAPDGRTLAPLPDHAEGLIIA
DINMEIAPAKAINDPVGHSKPEATRLVLDGHRPMTRVHSHSVIOEABEPHVOS
TAAPAVAVSOTQSDTLVQEPS"
BASE COUNT 343 a 558 c 476 g 350 t
ORIGIN

Query Match 63.2%; Score 1132.8; DB 1; Length 1727;
Best Local Similarity 99.8%; Pred. No. 8.9e-255;
Matches 1134; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 122 tatgcagacaagaanaatcgctccggcagccgcgcgtacagcgccgtcctcccaactacga 181

Db 303 TATGCAGACAAGAAAATATCTCCGGGACAGCCGCGTATACAGGCCGCTCTCCAACTACGA 362
QY 182 tctggaacagcggtgttgatataaaccaattgagctgcgtcgcgaagcccgatagaagcg 241
Db 363 TCTGGCAACGGGTGTGATATAAACCATTAAGCTGGCTCGTACAGGCCCGGCATAGAGGCTG 422
QY 242 tgaactgatcgtgtctgtgttaaaccttgagctgcgcgcgcgtacccctccacgctcgtgcg 301
Db 423 TGACCTGATGCTGTTTGTTGGTAAACCTGCGCCCGGCTATCCCTTCACAGCTCTGGCTGG 482
QY 302 cgcacgcgcctggctgcgtgaataacagltgcgcgtactatgcgaactgcgtcgtgcga 361
Db 483 CGACCGGCGCTGCTGCTGCTGAATAATACAGTGCCTCTACTATGACCACTCGCTCGGTGA 542
QY 362 caatgcaagattccaagcattgcgcaagcgaggaacttggatattcatgcacat 421
Db 543 CAGTGCAGAGTTTTCACAGCACTTGCCTCAGCGCAGGACCTTGGTATTTTCATTCGCACT 602
QY 422 gggctatagcagcagcagcgcgagcagccttacccttgagcgaattgcctgaatgcagaca 481
Db 603 GGGTTATAGCGAGCGCAGCGGGCGGCAAGCTTTACTGSGGCAATGCTGATGACGACAA 662
QY 482 gggcccaagatgctgtgtgcgtgcgaactcaaacctcaacatgttgagcgcgcgtgt 541
Db 663 GGGCCAGATGCTGTGTGCTGCGCGTGCGAACCACTACACATGTGAGCGCACCGTGT 722
QY 542 tggtagaaggtctgagcccggaatctgattgtctgcgaacacgagcctggccgcgtgcgtgc 601
Db 723 TGGTGAAGGTTTATGCCGAGATGTGATTTGTCTCGACACCGAGACTGGCGGCTCGGTGC 782
QY 602 cctgtgctgtgtagagacactgtccctctgagcaagtagcgcgtgtaactccagacga 661
Db 783 CTGTGCTGCTGGAGACACTGTCCCTTGACCAAGTACGCGCTACTCCACGACAGA 842
QY 662 agccattcaactatgcgcgcgcgcgcgtccttgcgtgtaaggaagaaagcccatgcgtc 721
Db 843 AGCCATTACATTGCGCGCTGCGCTCTTTGCTGTACAGGAACAGGCCATCGCTGCT 902
QY 722 caagcgaaggtgaacaatgctgctgcctgcgaactcattcgtgtgaagcgcaagctcttac 781
Db 903 CAGCGCCAGGTGATACATGCTGCTGCTGCAAACTATGCGTTGAAGGCCACGCTTTAC 962
QY 782 catgcgcgcagcagcgtgtgtcaccacgaagagacatgacatgctgtgaagtagtgaaca 841
Db 963 CATCGCGCGCAGCAGTGTCTGTACCCAGAGACATGGAATGCTGGAATAGTAAGTAACA 1022
QY 842 caagcgtccctctgtgaagtgaggcgagcaattcatgattttggccggagcagag 901
Db 1023 CAACGCTCTCCTGTGAAGTGGGCGGCGGACATTCATGATTTTGGCCGGACGGAGC 1082
QY 902 caattgctccctaccctgcacacagatgacgaagcgctatcatcttgccgcatctgaacat 961
Db 1083 CACATTGCTGCTCTACTCTGCGACACAGATGCGGAAGGCTGATATGCGCATGTGAACAT 1142
QY 962 ggaagaatattgccttcgcgaagcgatcaagacccctgtgtgacactactcaaacccga 1021
Db 1143 GGAAGAAATTCCTTGCGCCAAAGCGATCAACGACCCGTGTGGCCACTACTCCAAACCCGA 1202
QY 1022 ggcacacccgctgtgtacttgacacttgaggacccgttgagcccatcaccgcggatgataccaa 1081
Db 1203 GGCACCCCGTGTGCTGTGACCTGGGGCACCTGTGAGCCCATAGCTCGGCTGCAATTCGA 1262
QY 1082 aagcgtgacccaagaagaagcctccgagccgacgctgcaagtaagtaagcctgacgcctgcg 1141
Db 1263 AAGCGTATTCAGAGAAACACTCCGACGCGACGTCGAAGATACGGCTCGCGCTCGCG 1322
QY 1142 cgtcagccagactcagagactcgatacgtactcgtgtgcaagaacgctcctgaccccaaaa 1201
Db 1323 CGTACGACAGACTACAGACTCGGATACGTTACGTGTGACGAAGAACCTGCTCGACCCCAAAA 1382
QY 1202 gatgacaagcccggaacaaactgtccgggtctctgattcctctcgtggtcccgagctc 1257
|||||

Db 1383 GATGACAAGCGCCGGGCAAACTGCTCGGGCTTGATTCCTCTCGCTGCCGATCC 1438

RESULT 4

LOCUS E07629 1727 bp DNA linear PAT 29-SEP-1997

DEFINITION gDNA encoding nitrilase.

ACCESSION E07629

VERSION E07629.1 GI:2175764

KEYWORDS JP 1994153968-A/6.

SOURCE JP 1994153968-A/6.

ORGANISM Alcalligenes faecalis.

Alcalligenes faecalis

Bacteria; Proteobacteria; beta subdivision; Alcalligenaceae;

Alcalligenes.

1 (bases 1 to 1727)

REFERENCE

AUTHORS Endo,R., Yamada,H., Shimizu,A., Nagasawa,T. and Kobayashi,T.

TITLE PRODUCTION OF ORGANIC ACID BY TRANSFORMANT CONTAINING GENE DNA

JOURNAL CODING POLYPEPTIDE HAVING NITRILASE ACTIVITY

Patent: JP 1994153968-A 6 03-JUN-1994;

NITTO CHEM IND CO LTD

COMMENT OS Alcalligenes faecalis

PN JP 1994153968-A/6

PD 03-JUN-1994

PE 17-NOV-1992 JP 1992306663

PI ENDO RYUICHI, YAMADA HIDEAKI, SHIMIZU AKIRA, NAGASAWA TORU, PI

KOBAYASHI TATSUHIKO

PC C12P7/40,C12P7/42,C12P11/00,(C12P7/40,C12P1:05),(C12P7/42, PC

C12R1:05),

PC (C12P11/00,C12R1:05);

CC strandedness: Single;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

EH key Location/Qualifiers

FT source 1..1727

FT /organism='Alcalligenes faecalis' FT

FT /strain='JM3' /clone='pNUM20'

FT 5'UTR 1..303

FT CDS 304..1374

FT /product='nitrilase'

FT 3'UTR 1375..1727

FT RBS 293..297.

FEATURES

source Location/Qualifiers

1..1727

/organism="Alcalligenes faecalis"

/db_xref="taxon:511"

BASE COUNT 344 a 585 c 449 g 349 t

ORIGIN

Query Match 61.2%; Score 1097.6; DB 6; Length 1727;

Best Local Similarity 97.9%; Pred. No. 1.6e-246;

Matches 1112; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 122 tatgagaagaagaataatcgtccggcgacgcgttaacagccgcctctcccaactaga 181

Db 303 TATGAGACAAGAAAAATGCTCGGGCAGCCGCGCTACAGCGCGCTCTCCCACTACGA 362

Qy 182 tctgcaacgagtgctgtgataaaccattgagctgctcgtacagcccgagtgagagctg 241

Db 363 TCTGGCAACGGGTGTGTGATTAACAACATGAGCTGGCTCTCAAGCCCGCATGAGGGCTG 422

Qy 242 tgaactgctgtgttctgttgaaacctgtgcgcgagctatccctccacgtctgcgtg 301

Db 423 TGACCTCATCGTGTGGTTCACAACCTGGCTGCCGCGTATCCCTTCACAGTCTGGCTGG 482

Qy 302 cgcacagccctgtgtcgtgaaataacagtgccgcgtactatgcaactcgtctcgtgta 361

Db 483 CGCACCGGCTGTGGTGTGAATAACAGTGCCTCGCTACTATGCAACTGCTCTCGCTGGA 542

Qy 362 cagtcagagtttcaacgcatgtccacgagccgcacgacctgggtatttcacgcact 421

Db 543 CAGTGCAGAGTTTCAACGATTTGCCAGCCGCGACGAGCACTTGGGATTTATCATGCACAT 602

Qy 422 gggttataagcagcagcagcgcgcgcgccttaacctggccaattgctgatacagacaa 481

Db 603 GGGTTATAGCGAGCGCAGCGCGGCGAGCCTTACTCGGCGCAATGCGTATGACGACAA 662

Qy 482 gggcagatgctgt 541

Db 663 GGGCCAGATGCTGTGGTGTGCGTCGTCGCAAACTCAACCTTACACATGTGTGAGCGGCGTGT 722

Qy 542 tggtaaggltatgcgccgagatcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 601

Db 723 TGGTGAAGTTATGCCCGCAGATCTGATTTGTGCCAGACGAGCTGCGCGGTGGTGC 782

Qy 602 cctgt 661

Db 783 CCGTGTGCTGTGCGAGCAGCTGTCCCGCTTGAGCAAGTACGGCGCTGTACTCCAGACGA 842

Qy 662 agccattacatgtgcgcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 721

Db 843 AGCCATTCAATTGCGCGCTGTGCGCTGTCTTTTGTGTGTACACGCAACAGGCCATGCGCT 902

Qy 722 cagcgccaaagtgaacatgt 781

Db 903 CAGCGCCAAAGTGACATGTGCTGCTGCTGCAAACTTCTTGTGGTTGAAGGCCATGCTTTC 962

Qy 782 catgc 841

Db 963 CATCGCGCGCAGCAGT 1022

Qy 842 caagcgtccctgt 901

Db 1023 CAACGCTTCTGCTGTGAAGTGGGGCGGCGGAGTTTCCATGATTTTGGCCCGGAGGAGG 1082

Qy 902 cacattgt 961

Db 1083 CACATTGGCTCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142

Qy 962 ggaagaattgtctcttcgcgaagcgatcaacgacccgtgtgtgtgtgtgtgtgtgtgt 1021

Db 1143 GGAAGAAATTTGCTTGTGCGCAAGCGATCAACGACCTGTGGGCCACTACTCCAAACCCGA 1202

Qy 1022 ggcacccgt 1081

Db 1203 GGCACCGCTGT 1262

Qy 1082 aagcgtgatccagaagaagctcccgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1141

Db 1263 AAGCGTATCCAGCAAGAAGCTCCGAGCGGCACTGCAAACTACGCCCTGCCCGCTGCC 1322

Qy 1142 cgtcagcagactcagagctcgatagctactggtgtgaagaacgctctgaaccccaaa 1201

Db 1323 CCTACGCCACACTGTGAGTGCATGCGTACTGTGGTGAAGAAGCGTCTGTGACCCCAAA 1382

Qy 1202 gtlgacaagcccgagcaaacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1257

Db 1383 GATGACAAGCGCCGGGCAAACTCTCCCGCTTATCTCTGTGCTGCCGATCC 1438

RESULT 5

LOCUS E07629 7508 bp DNA linear BCT 15-APR-1994

DEFINITION E.coli RNA operon (rnb) coding for Glu-tRNA-2, 5S, 16S and 23S

ACCESSION J01695

VERSION J01695.1 GI:147581

KEYWORDS

SOURCE Escherichia coli DNA.

ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

Escherichia.

REFERENCE 1 (bases 1518 to 3059)

AUTHORS	Brosius,J., Palmer,M.L., Kennedy,P.J. and Noller,H.F.
TITLE	Complete nucleotide sequence of a 16S ribosomal RNA gene from <i>Escherichia coli</i>
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 75, 4801-4805 (1978)
REFERENCE	
AUTHORS	2 (bases 811 to 1543)
TITLE	Georadas-Tsch.E., Boros,I. and Venetianer,P.
JOURNAL	Structure of the promoter region for the rnb gene in <i>Escherichia coli</i>
REFERENCE	Nucleic Acids Res. 7, 2189-2197 (1979)
AUTHORS	80101075
TITLE	3 (bases 3500 to 6403)
JOURNAL	Brosius,J., Dull,T.J. and Noller,H.F.
REFERENCE	Complete nucleotide sequence of a 23S ribosomal RNA gene from <i>Escherichia coli</i>
AUTHORS	Proc. Natl. Acad. Sci. U.S.A. 77, 201-204 (1980)
TITLE	80145594
JOURNAL	4 (bases 1 to 7508)
REFERENCE	Brosius,J., Dull,T.J., Sleeter,D.D. and Noller,H.F.
AUTHORS	Gene organization and primary structure of a ribosomal RNA operon from <i>Escherichia coli</i>
TITLE	J. Mol. Biol. 148, 107-127 (1981)
JOURNAL	82055570
REFERENCE	5 (bases 6412 to 6507; 6603 to 6657)
AUTHORS	Singh,B. and Apirion,D.
TITLE	Primary and secondary structure in a precursor of 5S rRNA
JOURNAL	Biochim. Biophys. Acta 698, 252-259 (1982)
REFERENCE	83049084
AUTHORS	6 (bases 6518 to 6931)
TITLE	Sarmientos,P., Sylvestre,J.E., Contente,S. and Cashel,M.
JOURNAL	different stringent control of the tandem e. coli ribosomal rna promoters from the rna operon expressed in vivo in multicopy plasmids
REFERENCE	Cell 32, 1337-1346 (1983)
AUTHORS	83180429
TITLE	7 (bases 3380 to 3509; 6392 to 6503)
JOURNAL	King,T.C., Sirdeshmukh,R. and Schlessinger,D.
REFERENCE	RNase III cleavage is obligate for maturation but not for function of <i>Escherichia coli</i> pre-23S rRNA
AUTHORS	Proc. Natl. Acad. Sci. U.S.A. 81, 185-188 (1984)
TITLE	84119433
JOURNAL	8 (sites)
REFERENCE	Orosz,A., Boros,I. and Venetianer,P.
AUTHORS	Analysis of the complex transcription termination region of the <i>Escherichia coli</i> rnb gene
TITLE	Eur. J. Biochem. 201, 653-659 (1991)
JOURNAL	92037621
REFERENCE	9 (sites)
AUTHORS	Prestle,J., Schoenfelder,M., Adam,G. and Mundry,K.W.
TITLE	Type I ribosome-inactivating proteins dephurinate plant 25S rRNA without species specificity
JOURNAL	Nucleic Acids Res. 20, 3179-3182 (1992)
REFERENCE	92319651
AUTHORS	
TITLE	Location/Qualifiers
JOURNAL	1. 7508
REFERENCE	/organism="Escherichia coli"
AUTHORS	/db_xref="taxon:562"
TITLE	1. 250
JOURNAL	/note="putative VECTOR sequence lambda; putative"
REFERENCE	1068. 1072
AUTHORS	/citation=[2]
TITLE	1212. 1218
JOURNAL	/note="Promoter: RNA polymerase binding site"
REFERENCE	1223. >3306
AUTHORS	/note="rnb precursor RNA (alt.)"
TITLE	1331. 1337
JOURNAL	/note="alternate promoter"
REFERENCE	1344. >3306
AUTHORS	/note="rnb precursor RNA (alt.)"
TITLE	1507
JOURNAL	/citation=[2]
REFERENCE	/replace=""
AUTHORS	1518. . 3059
TITLE	rRNA

Query Match	27.9%	Score 500	DB 1	Length 7508
Best Local Similarity	97.1%	Pred. No. 1.6e-106		
Matches 509	Conservative	0	Mismatches 15	Indels 0
Db 1267	agaagtcgacgttcgaagcagcaagcgttggctccaccctgaccacatgctcgaactcagaag	1326		
Db 6490	ACAATTTCGCTTGCGGCGGCACTACCGCGGTGTCCACCTACCCCATGCCGAACCTGGAAG	6549		
Qy 1327	tgaaacgcgcgtatgcgcgaatggtagttgttgggtctccccaatcgagagtaaggaaatgccc	1386		
Db 6550	TGAAGACGCGGTAGCGCCGATGTAGTGGGTCTCCCATGAGAGATGAGAACTGCC	6609		
Qy 1387	agggatccaataaagcaagaaagctcaatgttgaagaagcttggtcttgcgtttatctgtgt	1446		
Db 6610	AGGCATCAATATAACGAAAGGCTCAGTGAAGACTGGGCGCTTTCGTTTATCTGTTGT	6669		
Qy 1447	tgtcgtgttaagcgtctcccttgatgtagaacaatccgcgcgagacgtatttgaacgttcgc	1506		
Db 6670	TTGTGCGTGAACGCTCTCTCTGATGATAGCAAAATCCCGCGGAGCGGATTTGAACGTGGG	6729		
Qy 1507	aagcaacgccccggaaggttgcgcgaggaagcgcgcgaataactgcaagcatcaaat	1566		
Db 6730	AAGCAACGCGCCGAGGCGTGGCGGAGACGCCGCCCATTAACCTCCAGCATCAAAAT	6789		
Qy 1567	aagcagaagagccatctctgaacgcatggtcccttttggtttctcaaaacttctgtgcgc	1626		
Db 6790	AAGCAGAAAGGCTCTGACGATGGCTTTTGGCTTCTACAACTTCTCTGTCTGTC	6849		
Qy 1627	atatacacaagcatcccccacagatacgttaactagccctgttttgatcagaagaa	1686		
Db 6850	ATATCTACAGCATTCGCCCCACAGATACGATTAACCTAGCTCGTTTTCATCAGGAAA	6909		
Qy 1687	gcagcatgaacacactcctttaaacccttgaacacatlttgcatlgtatcataatgctcag	1746		
Db 6910	GCAGCATGAACACACTCCTTAACCCCTGGAACACATTTGGCATTCATATGCTCAG	6969		
Qy 1747	caactgtatgtgcgcgaagaagcaacaacaattractcaatgctccg	1790		
Db 6970	CACATTTGATGTGCCGAAAGCAGAACACAACTTACTCAATGACCTG	7013		

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
6496..6615
/note="5S rRNA (rnbB)"
rRNA

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
6496..6615
/note="5S rRNA (rnbB)"
rRNA

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
6496..6615
/note="5S rRNA (rnbB)"
rRNA

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
6496..6615
/note="5S rRNA (rnbB)"
rRNA

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
6496..6615
/note="5S rRNA (rnbB)"
rRNA

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
6496..6615
/note="5S rRNA (rnbB)"
rRNA

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
6496..6615
/note="5S rRNA (rnbB)"
rRNA

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
6496..6615
/note="5S rRNA (rnbB)"
rRNA

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
6496..6615
/note="5S rRNA (rnbB)"
rRNA

BASE COUNT 1969 a 1694 c 2166 g 1679 t

ORIGIN 1 bp upstream from BamHI site.

Query Match 27.9% Score 500; DB 1; Length 7508;
Best Local Similarity 97.1%; Pred. No. 1.6e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0

trNA /note="16S rRNA (rnbB)"
3231..3306
/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"
/product="rRNA-Glu"
3500..6403
rRNA /note="23S rRNA (rnbB)"
649

JOURNAL REFERENCE AUTHORS	82055570 2 (bases 1 to 7508) van Keulen,H., Guteil,R.R., Gates,M.A., Campbell,S.R., Erlandsen,S.L., Jarrold,E.L., Kulda,J. and Meyer,E.A. Unique phylogenetic position of Diplomonadida based on the complete G. diodenalis and Hexamita sp rRNA sequence of Giardia ardeae, G. muris, FASER J. 7 (1), 223-231 (1993)
JOURNAL MEDLINE REFERENCE AUTHORS	93138311 3 (bases 1 to 7508) Weltzmann,C.J., Cunningham,P.R., Nurse,K. and Ofengand,J. Chemical evidence for domain assembly of the Escherichia coli 30S ribosome FASER J. 7 (1), 177-180 (1993)
JOURNAL MEDLINE AUTHORS	93138304 4 (bases 1 to 7508) Ringquist,S., Cunningham,P., Weltzmann,C., Foremenoy,L., Pleij,C., Ofengand,J. and Gold,L. Translation initiation complex formation with 30 S ribosomal particles mutated at conserved positions in the 3'-minor domain of 16 S RNA J. Mol. Biol. 234 (1), 14-27 (1993)
JOURNAL MEDLINE REFERENCE AUTHORS	94047056 5 (bases 1 to 7508) Sako,Y., Takai,K., Ishida,Y., Uchida,A. and Katayama,Y. Rhodochromus obomensis sp. nov., a modern lineage of extremely thermophilic marine bacteria Int. J. Syst. Bacteriol. 46 (4), 1099-1104 (1996)
JOURNAL MEDLINE COMMENT FEATURES	97016816 On May 8, 1997 this sequence version replaced gi:42877. Location/Qualifiers 1..7508 /organism="Escherichia coli" /db_xref="taxon:562" 240..254 /note="lambda attachment site" 275..1144 /note="unidentified reading frame I" /codon_start=1 /transl_table=1 /protein_id="CAA23638.1" /db_xref="GI:42878" /db_xref="SWISS-PROT:P22634" /translation="MROSMATKLDGDGTPTCLAIATPSPEPRPLVLEDSVGGLSYDEIT RLHLPLDHYIYAEENVAFPYGEKSEAFIVERVAIVAVGEREYLAAVACNASTVS SLPLRKEPDFEPYGVVPAPKIPARLRATLVGLATRGVRKYTHELLARFANECQ LEMGSAMVELAELAKHGEDVSLDALRIKRPVLRMKPEPDVTVLGGTHPPILOEEI LQVPEGRVLVSGCAALARTARTAWLHEHADPAKSADANIINFCMMNTGCAGDLPLVDLR YGPETLEKLAYLG"
rRNA	1518..2959 /product="16S ribosomal RNA"
tRNA	3231..3306 /note="transfer RNA - Glu - 2"
rRNA	3500..6403 /product="23S ribosomal RNA"
BASE COUNT	1970 a 1694 c 2166 g 1678 t
ORIGIN	
Query Match	27.9% Score 500, DB 1; Length 7508;
Best Local Similarity	97.1%; Pred. No. 1,6e-106;
Matches 509; Conservative	0; Mismatches 15; Indels 0; Gaps 0;
Oy 1267	agagcgcacccgcaaggcatgtcaagcttgggtcccaccgcgaacctatgccgaactcagaag 1326
Db 6490	AGAATTTCCTCGCGCCGACGTAACCGGTGTCACCACCTGACCCCCTGCCGAACCTAGAG 6549
Oy 1327	tgaaagcccgtagcgcacgatgtagtgtggygggtctccccatgcygagtaggaaactgcc 1386
Db 6550	TGAAAACGCCGTAAGCAGCCCATGTGTGGGTCTCCCCATGCGAGACTAGGAACCTGCC 6609
Oy 1387	agggcataataaaaagaaggtctcagtcgaaagaactgycggcttcattatctgttgt 1446

Db	Accession	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT				
Db	6610	AE000471	Escherichia coli K12 MG1655 section 361 of 400 of the complete genome.	AE000471	U000096	AE000471.1	GI:1790404	Escherichia coli K12. Escherichia coli K12 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.	1	(bases 1 to 11679)	Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.	The complete genome sequence of Escherichia coli K-12	Science	277 (5331), 1453-1474 (1997)	97426617	9278503	2 (bases 1 to 11679)	Blattner, F.R.	Direct Submission	Submitted (16-JAN-1997)	Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459
Db	6670	AE000471	Escherichia coli K12 MG1655 section 361 of 400 of the complete genome.	AE000471	U000096	AE000471.1	GI:1790404	Escherichia coli K12. Escherichia coli K12 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.	3	(bases 1 to 11679)	Blattner, F.R.	Direct Submission	Submitted (02-SEP-1997)	Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459							
Db	6730	AE000471	Escherichia coli K12 MG1655 section 361 of 400 of the complete genome.	AE000471	U000096	AE000471.1	GI:1790404	Escherichia coli K12. Escherichia coli K12 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.	4	(bases 1 to 11679)	Plunkett, G. III.	Direct Submission	Submitted (13-OCT-1998)	Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA							
Db	6850	AE000471	Escherichia coli K12 MG1655 section 361 of 400 of the complete genome.	AE000471	U000096	AE000471.1	GI:1790404	Escherichia coli K12. Escherichia coli K12 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.	5	(bases 1 to 11679)	Plunkett, G. III.	Direct Submission	Submitted (13-OCT-1998)	Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA							
Db	6910	AE000471	Escherichia coli K12 MG1655 section 361 of 400 of the complete genome.	AE000471	U000096	AE000471.1	GI:1790404	Escherichia coli K12. Escherichia coli K12 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.	6	(bases 1 to 11679)	Plunkett, G. III.	Direct Submission	Submitted (13-OCT-1998)	Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA							
Db	6970	AE000471	Escherichia coli K12 MG1655 section 361 of 400 of the complete genome.	AE000471	U000096	AE000471.1	GI:1790404	Escherichia coli K12. Escherichia coli K12 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.	7	(bases 1 to 11679)	Plunkett, G. III.	Direct Submission	Submitted (13-OCT-1998)	Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA							


```

misc_feature
1..1257
/gene="misc_feature"
<1..1737
/note="corresponds to M21516; ECOKATGA(1225..2805)"
<1..1257
/gene="katg"
/standard_name="catalase HPI; catalase-peroxidase; catalase"
/EC_number="1.11.1.6"
/note="CG Site No. 14983"
/codon_start=1
/transl_table=11
/product="catalase hydroperoxidase I"
/protein_id="AAC43048.1"
/db_xref="GI:396289"
/translacion="AITSGLLEVWVTQTPTQNSNFENLEFKYEMVQTRSAPAGAIQFEAF
VDAPILIDPDPSPSKRRPTMLVLDLTLRPPEEKKISRRLNPOAENEFARXWFE
LTHRMGKRSRYIGPEVRKEDLIQDDPLPQIYNPEODIIDLFAIDSLASVEL
SVAMASCTPFRGDKRGANGARLAIPOBMDVAAVRAALPVLEKIQKSGASGLS
DIIVAGVGVGVKASASAGSLIHVPFARGVARDARODOTDIEMFELPEIADGPNRYR
RLDVSTTSLSLIDKRAQQLTTPAPEKALTYVGGRRVLGAFNFDSSKNVFTLDRGVLSNDF
FVNLDIMDEWMAIDESKELFEGRDRETEGVKFTASRADLVFGSNSVLRAAVEAYASS
DAHEFVMDYFAAAMVKVNLDRFDLL"
1..93
/gene="katg"
/note="93 bp overlap with end of L19201 (ECOMU87)"
<1..161
/gene="katg"

```

```
misc_difference /note="corresponds to lambda clone EC14-54" 938
                /gene="katG"
terminator      /note="G in M21516; C here" 1266..1305
                /note="putative" 1317..2255
CDS             /note="similar to Desulfurolobus ambivalens hypoch. 28.3
```

```
/codon_start=1
/transl_table=11
```

```

/protein_id="AAC43049.1"
/db_xref="GI:409786"
/translation="MGERRLHTGSLMSAAGKSNPLAISGLVLLITNSYSIPIFKQVY
SYIGAFVETALRCIIPGALVLEIVLLRERGRPFPEKVTIALIALQETGAWGLAKOMALA
VSGGKFAVLTSTYMPFVWVFEALFELRERIRGQYFALIAEFLFVLPQOLFDEGK
SMKSMALLISLGSMSGASAIYAKRLIAHPHYDLSTLSQWLVAAVMSVALLPQ

```

```
misc_difference
1429, .1432
/note="TTT in M21516; TTTT here"
misc_difference
1720
/note="C in M21516; G here"
```

```

CDS
    complement(248..2899)
        /codon_start=1
        /transl_table=11
        /label=ORF_f205
        /protein_id="AAC43050.1"
        /db_xref="GI:396291"
        /translation="MKASLALSLTAFSTHSIKSPVAPPTVQOANTLADGAF
        /COIGSLTIDPAYDLITPGGDVPEKRCVCSVYIALRSKVDLQKLVYEDMAKNF
        EYPQWKLKIDPDSNIDHRRVNPILTEFSEHDKTRTSPNSPDYQAGDITVSRNLNGLA

```

```

promoter      HIGVYSDGFARDGTELVTHNIGAGAGQEDVLFNMNMVGHYRFPVK"
               complement(2886..2915)
misc_structure /note="promoter matrix score of 42; putative: within ORF
               205,"which would suggest alternate start codon"
               2912..3053
               /note="predicted bend of 81 degrees"
repeat_region 3031..3064
               /standard_name="REP; repetitive extragenic palindromic
               element"
               /note="contains 1 REP sequence"
misc_structure 3090..3219

```

```
gene /note="predicted bend of 75 degrees"
      complement(1174..4316)
      /gene="glda"
CDS complement(1174..4316)
      /gene="glda"
      /note="similar to Bacillus stearothermophilus glycerol
      dehydrogenase"
      /codon_start=1
      /transl_table=11
      /label_ORF_f380
      /product="glycerol dehydrogenase"
      /protein_id="AAC43051.1"
      /db_xref="GI:396292"
      /translation="MPHLLLSKGAIMDRILIOSPKYIOGADVIRLGEYLKPLXER
      MLVGDKFPYGAOSTVEKSEFKDAGIVYIAPGEGSCONEIDRLGIAETAGCAAIL
      GIGGSKTIDTAKALAHFMGVPAIAPTASTAPCSALSVITDEEPRIRILLRNP
      NMVIVDTKIVAGAPARLNAIGDALATWEFARACSRGATTAGGCTOALALAEI
      CYNLTLEEGEKAMLAEOHVTPALERVLEANTYLSGVFEGSGLAHAHANGILTAI
      PDHHYHGEKVAFGTLTQVLENAPEVEITVAALSHAVGLPTLAOLDIKEDVPK
      MRIVAAACACEGETIHNMPGATPDQVYVALLVADQYGGRFIOEME"
misc_structure 4022..4156
      /note="predicted bend of 75 degrees"
      complement(4288..4950)
CDS /note="similar to Bacillus subtilis hypoth. 20 kDa
      protein, in tsr 3' region"
      /codon_start=1
      /transl_table=11
      /label_ORF_f220
      /protein_id="AAC43052.1"
      /db_xref="GI:396293"
      /translation="MELYLDTANVAEVERLARIPIAGVTNPSSIIAASKSEIMEVLP
      RLOKAIIGDEGILFAOTMSRDAGOWEAEARLDAIPGIYVKIIPVSEGAALKILKE
      GTTGTCTAYSAAGCLALACAGKAYAPRVNVADOGGCTITVOELTLLMHAPES
      MYLAASFKTPRALDCLACESTITLPDVAQOMLTPAVESIEKEFHDMMARFTT
      HL"
promoter complement(4922..4949)
      /note="promoter matrix score of 47; putative; within ORF
      f220, which would suggest alternate start codon"
CDS complement(4962..7097)
      /note="similar to phosphotransferase system enzyme I"
      /codon_start=1
      /transl_table=11
      /label_ORF_f711
      /protein_id="AAC43053.1"
      /db_xref="GI:409787"
      /translation="WCSGSAGGILTPISLIDNALGNLPAKGVDAEQSALENLTV
      LKNIERLSDSGATSALIEAHRSLAGDTSLEBHLAAGSAGSCAEALVASANHFCE
      EFSRSSSYLRALDVRVCFOLQOITGEORFPAPGKLTPOALCMADELTPSOELE
      LDKNHLKGLLSGGTSHVILARSFNIPLVGVVIDALTPMOOCTIYIDNAGATV
      VERGEAVARYOQAEARVODALREORVMTTOCARADGRIRIETIANIASHVQAQAF
      NCAGEVGLFRTMLYMDRTSAGESELYNIFCOALBSANGRSIYATMDIGDKPVYD
      LNIPEANPFLIGRAVRIEYASLETTQRLSILKASAHGSLKIMIPMISMEELIW
      KEKLAQAQOLRNEHLPDEKIQGLIMLEPSYMEIIDCCCEIDFESIGSNDLYL
      LAVDRNNAKVTRHYNLSNPAFLDALYAAQVAVHOGKIGLCELGAKGSVLPILVGL
      GLDELMSAPSIIPAKARMAOLDISRECRLLNOAMACRPSLEVEHLLAOFRTQODAP
      LYTAFCTLESMDRSKEVILKGMTDLILAGRCYRKLLEADILAEAVFSTGICPSF
      AIPHSKSEHIEOSTISVARLOAPVRKGDDEAOPITMLTINKHAAGQHNRIISRILAR
      IMHEERNALVNAASADAIASLQHELEL"
CDS 7099..7428
      /codon_start=1
      /transl_table=11
      /label_ORF_o109
      /protein_id="AAC43054.1"
      /db_xref="GI:409788"
      /translation="MRARIICGFRLVSETGSSSSSLLTSASGASGQMSRSRHLKR
      WCASCSAPEISNMOLSPASVSPISARALLALPCGCAGATQFWMKKRYTAFORGL
      PDAHYRR"
promoter complement(7108..7136)
      /note="promoter matrix score of 46; putative"
misc_feature 7200..72839
      /note="corresponds to lambda clone EC21-52"
terminator 7578..7699
      /note="putative"
```

```
promoter 7771..7800
      /note="promoter matrix score of 48; putative; within ORF
      0359, which would suggest alternate start codon"
CDS 7771..8850
      /note="similar to phosphotransferase system enzyme II"
      /codon_start=1
      /transl_table=11
      /label_ORF_o359
      /protein_id="AAC43055.1"
      /db_xref="GI:396296"
      /translation="MNEIVQILKNTROHLMTGVSHNIPFVSGGILLAVSVLYKGA
      VPDVAADPNLKRFLDVGALTLMPFLAAVYIGYSIAERSALAPALGAWGNSFGAG

Query Match      27.9%; Score 500; DB 1; Length 176195;
Best Local Similarity 97.1%; Pred. No. 1.9e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1267 agagtcgacctgcagcagcatgcaagcttgagtlccaccctgaccccatgccaagcagaag 1326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36873 AGAATTTTGCTGGCGGAGTAGAGCGGGTGCTCCACCTGACCCCATGCCAGACTCAGAAG 36932

QY 1327 tgaacgcgctagcgcgcagatgtagtgggtcctcccatgcgagagtaggaactgccc 1386
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36933 TGAACGCCCTAGCGCGGATGATGTCGGGCTCCCTCCATGGCAGTAGGGAACCTGCC 36992

QY 1387 aggcattcaataaagaagaagctcagtcgaagaagcttgagcctttttatcttctt 1446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36993 AGGCATCAATAATAAGAAAGGCTCAAGTCAAGAAAGCTGGGCTTTGCTTTATCTGTGT 37052

QY 1447 ttgcgtgtagcagctcctcctgtagtagacaatacgcgcggagacgatttgaacgttgcg 1506
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37053 TTGTCGGTGAACGCTCTCCAGTAGACAAATCCGCGGAGCGGATTTGAACGTTGCG 37112

QY 1507 aagcaacgcgcgcggaggtgtaggcgagacgcgcgcacataaacctgcagacatcaaat 1566
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37113 AACCAACGCGCCCGAGGAGGTGGCGGCAAGACGCCGCATTAACCTGCCAAGCATCAAT 37172

QY 1567 aagcagaagccatccctcgaagatgagccttttggcgtttcctaaactctctgctgc 1626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37173 AACCAAGGCGCATCTCGACGATGGCTTTTGGCTTTCAAACTCTCTGCTGCTG 37232

QY 1627 atatctacaagcattccccacacagatacagtaaacatgacctgcttttgcacagaaa 1686
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37233 ATATCTCAAGCAGCATCCGCCACAGATAGCGTAACTAGCTGTTTTCATCAGAGAA 37292

QY 1687 gcsagctatgaacacactccttaaaacccctggaacacattggcatgtacataatgctcag 1746
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37293 GCAGCTATGAACCACTCTTAAACCTTGGAACACATTTGGCATTTGATCATATGCTCAG 37352

QY 1747 caaatgtatgtgtccgaagcagaacaataactactaatgccc 1790
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37353 CAAATGTATGTGTCCGAAGCAACCAATTTCTCAATGCTGTG 37396

RESULT 10
SYNCGNABP LOCUS SYNCGNABP 860 bp DNA linear SYN 27-APR-1993
DEFINITION e.coli rna operon promoters fused to rnb operon terminators.
ACCESSION K00764
VERSION K00764.1 GI:209262
KEYWORDS 16S ribosomal RNA; 5S ribosomal RNA; promoter region; ribosomal
RNA; rna operon; rnb operon; terminator.
SOURCE Escherichia coli dna; clones psl and p55.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 860)
AUTHORS Sarmientos,P., Sylvester,J.E., Contente,S. and Cashel,M.
TITLE differential stringent control of the tandem e. coli ribosomal rna
promoters from the rna operon expressed in vivo in multicopy
plasmids
JOURNAL Cell 32, 1337-1346 (1983)
MEDLINE 83180429
COMMENT [1] assessed stringent control of the p1 and p2 promoted
```

transcripts *in vivo*. It was found that the upstream (p1) promoter was subject to stringent control, while the downstream (p2) promoter activity was inhibited by amino acid starvation in both stringent and relaxed hosts. A plasmid with an additional deletion of the p2 region also showed stringent regulation of the p1 promoter. *In vivo* as well as *in vitro* transcripts arising from both promoters terminate predominantly in the t1 terminator region about 40 bp beyond the mature rnb 5s rna gene. The exact termination position was ambiguous in [1]. Termination region t2 spans bases 723-751. The rnb terminators t1 and t2 were deduced from sequence homologues with known terminators. The rna promoter sites were verified *in vivo*. The first 40 bases of the mature 5s rna are missing in the fused product. clone p55 contains two additional 'c' residues at positions 158, 159 that were contributed by hindIII linkers.

FEATURES		location/qualifiers
source		1..860
		/organism="synthetic construct"
		/db_xref="taxon:32630"
misc_RNA		69..570
		/note="16S precursor RNA p1"
misc_RNA		186..446
		/note="5S precursor RNA p2"
rRNA		361..446
		/product="16S ribosomal RNA"
misc_RNA		447..570
		/note="1 precursor RNA"
rRNA		447..544
		/product="5S ribosomal RNA"
BASE COUNT	243 a	215 c
ORIGIN	374 bp upstream of saulilla site.	214 g 188 t

Query Match	Similarity	23.5%	Score 421	DB 12	Length 860
Best Local	Similarity	100.0%	Pred. NO. 4.8e-88		
Matches 421	Conservative	0	Mismatches	0	Indels
					Gaps
QY 1288	aagcttggctcccaactgacccatgcccgaactcagaagtgaacgcgctagcgcgcatg				
Db 440	AAGCTTGGGTCCACCTGACCCCATGCCGAATCTCAAGAACTGAACGGCCGTAGCCGGATG				
QY 1348	gtatgctgggggtctcccatgcaagatgaggaaatgccaagcatcaaatlaaaacgaag				
Db 500	GTAAGTGGGGGTCTCCCATGCGAAGTAGGGAAGTGCAGGCATCAATAATTAACCAAG				
QY 1408	gtctacgtcgaagaagacttggcctcttcgttttactctgttctgtctgtcgtgaagctctcgt				
Db 560	GCTCAGTGTGAAAGACTGGGGCTTTGCTTTTAACTGTTGTTTGTGCGGTGAACCTCTCTG				
QY 1468	agtaagacaacaatccgcgcgggagcgaatttgaaacgttgcgaacaacagcgccggaggtgtg				
Db 620	AGTAGGACAAATCCCGCGGAGCGGATTTGAACGTGGCAACACACGGCCGGAGGTGG				
QY 1528	cgggacgaagcgcgcgcacaaacttgcagggcacaatcaagcagaagggcactctctgacg				
Db 680	CGGCGACGAGCGCCCGCATTAACCTCCAGGCATCAAAATTAAAGCAGAGGCATCTGTAGC				
QY 1588	gatggccttttgggttctctaaacactctctgttcgtcatatctcaaaagccatccccc				
Db 740	GATGGCCTTTTGGCTTTCTTACAAACTCTCTCTGTCTATTTCTTCAACAGCATCCCCCC				
QY 1648	acagaatacgttaaacatacgtcgtcttcttgcacacagaagaacgacataatgaacactctta				
Db 800	ACAGATACGGTAACTAGCTCTGTTTTCGATCAGAGAAAGCAGCTATGACCACTCTTAA				
QY 1708	a 1708				
Db 860	A 860				

DEFINITION	pKK23-8 cloning vector, complete sequence.
ACCESSION	U13859
VERSION	U13859.1 GI:595745
KEYWORDS	chloramphenicol acetyltransferase; promoter analysis; beta-lactamase.
SOURCE	unidentified cloning vector. unidentified cloning vector artificial sequence; vectors.
ORGANISM	1 (bases 1 to 5094)
REFERENCE	Malone,J.A.
AUTHORS	pKK23-8: A prokaryotic promoter analysis vector featuring a promoterless chloramphenicol acetyltransferase gene for quantitative analysis of promoter strength
TITLE	Unpublished (1994)
JOURNAL	2 (bases 1 to 5094)
REFERENCE	Brosius,J.
AUTHORS	Plasmid vectors for the selection of promoters
TITLE	Gene 27 (2), 151-160 (1984)
JOURNAL	84209845
MEDLINE	3 (bases 15 to 58; 1243 to 1568; 1741 to 2066)
REFERENCE	Brosius,J., Dull,T.J., Sletter,D.D. and Noller,H.F.
AUTHORS	Gene organization and primary structure of a ribosomal RNA operon from Escherichia coli
TITLE	J. Mol. Biol. 148 (2), 107-127 (1981)
JOURNAL	82055570
MEDLINE	4 (bases 1 to 5094)
REFERENCE	Malone,J.A.
AUTHORS	Direct Submission
TITLE	Submitted (19-AUG-1994) James A. Malone, International Technical

```

terminator      2039..2066
                /standard_name="rrn T2"
                /citation={2}
                complement(3267)
rep_origin      /note="base 3267 represents the first base of the newly
                synthesized strand"
                /direction-left
                complement(4026..4956)
gene            /gene="bla"
                complement(4026..4886)
CDS             /gene="bla"
                /codon_start=1
                /transl_table=11
                /product="beta-lactamase"
                /protein_id="AAC53597.1"
                /db_xref="GI:595747"
                /translation="MSIQHFRVALIPFFAFCFLPVFAHPETLVKVKAAEDLQARNGV
                IELDINSKGLIESEFPERFPMSTFKVLGAVLSRYDAGQEQGRIRIHSQNDLVE
                YSPYERKHLTDGMTVRELCSAATMTMSDNTANLLITIGRKELTAPLPHNGDHYTRL
                DRWEPELNEAIPNDERDPTMPAAMATLRKILITGELTLASRQOLIDMWEADKVGPL
                LRSALPACMFIAADSGAGERSRGIIAALGPDGKSRIVITYTTSQATMDERNQIA
                ELGASLIRHW"
                complement(4928..4956)
promoter        /gene="bla"
BASE COUNT     1293 a 1290 c 1290 g 1221 t
ORIGIN
Query Match      22.5%; Score 404; DB 12; Length 5094;
Best Local Similarity 96.5%; Pred. No. 5e-84;
Matches 413; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1267 aagatcgacctgcagcgatgcgaagcttggtcccaactgaacccatgccgaactcagaag 1326
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 1237 AGAATTTGCTGGCGGCACTAGCGCGGTGTCACCACTGACCCCATGCCAATCTCAGAAG 1296

QY 1327 tgaacgcgcgtagcgccgcatggtgtagtgtggtgtcccatcgagaagtaggaaactgcc 1386
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 1297 TGAACGCCGCTAGCGCGCATGTAGTGTGGGTCTCCCATCGAGAGTAGGGAATGCC 1356

QY 1387 aggcatacaataaagaaagcctcagtcgaaagactgggaccttcgtttatctgtgtc 1446
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 1357 AGGCATCAATTAAGAAAGAAAGCTCAGTGAAGAAAGCTGGGCTTTGCTTTATCTGTGT 1416

QY 1447 ttgtggtgaagcgtctctctgtgtagtgaacaatccgcggagcgagattgaacgttgcg 1506
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 1417 TTGTGGGTGAACGCTCTCTGAGTAGACAAATCCCGCGGAGCGGATTTGAACGTTGCG 1476

QY 1507 aagcaacgcgcgcggaaggtgtagcggaagcgcgcgaataaactgcaggaatcaaatc 1566
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 1477 AAGCAACGCCGCGGAGAGGTGGCGGAGAGACCGCCGCATTAACCTCCAGGCATCAAAAT 1536

QY 1567 aagcagaagcgaatccctgaagatggccttttcgctttctacaacatctcttcgtcgtc 1626
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 1537 AAGCAGAAGGCATCTCGACGATGGCTTTTGGCTTTCTACAAACCTCTCCGTCGCTC 1596

QY 1627 atatctacaagcatccccccacagataagtgtaaaactaagcttcgttttcgaatcaga 1686
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 1597 ATATCTACAAAGCATCCGCCACAGATAGGTAACACTGCTCGTTTTCATCAGAGAAA 1656

QY 1687 gcaagctat 1694
    ||| | ||| |
DB 1657 GCAGCTGT 1664

```

```

RESULT 12
AF216802/c 6733 bp DNA linear SYN 22-JUN-2000
LOCUS AF216802 Shuttle vector pDL278, complete sequence.
DEFINITION AF216802
ACCESSION AF216802.1 GI:8650401
VERSION AF216802.1
KEYWORDS Shuttle vector pDL278.
SOURCE

```

```

ORGANISM Shuttle vector pDL278
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 6733)
AUTHORS Leblanc,D.J., Lee,L.N. and Abu-Al-Jaibat,A.
TITLE Molecular, genetic, and functional analysis of the basic replicon
of pVA380-1, a plasmid of oral streptococcal origin
JOURNAL Plasmid 28 (2), 130-145 (1992)
MEDLINE 93028887
PUBMED 1409970
REFERENCE 2 (bases 1 to 6733)
AUTHORS Dunny,G.M. and Bae,T.
TITLE Direct Submission
JOURNAL Submitted (15-DEC-1999) Microbiology, University of Minnesota, Box
196 UMC, 1460 Mayo Building, 420 Delaware St. SE, Minneapolis, MN
55455, USA

FEATURES
source
Location/Qualifiers
1..6733
/organism="Shuttle vector pDL278"
/db_xref="taxon:111473"
/focus="Gram positive-Gram negative shuttle vector"
/note="Gram positive-Gram negative shuttle vector"
1682..2449
/codon_start=1
/transl_table=11
/product="spectinomycin adenyltransferase"
/protein_id="AAF78194.1"
/db_xref="GI:8650402"
/transl_table="MRIRYLMVYEQINKVKILRKHLNNLIGTYMGSGVSGSLKPN
SDLFLVYVSPRLDQSKELIOLKIRPSKIGDSNRYIELTIIIOEVMYPMNHP
KQEFYGEMLQELVQGYIPOKELNSDLITIMLYQAKRNKRIRYGDLEELDIPFS
DVRRAIMDSSEELIDNYODDENSTLTICRMILPMDDKIIPKDIAGNAVAESSPLEH
RERLIAVRSYLGEMTENVNLTITVYLNLRKLK"
complement(4639..4878)
/note="orf 1"
/codon_start=1
/transl_table=11
/product="unknown protein"
/protein_id="AAF78195.1"
/db_xref="GI:8650403"
/transl_table="WSEKELAIIPKKKVFSGKEYVELTKEDPBEFKGLIYSRNLWHO
KELENEQLRIYVPLRKRFRFASWNLKKVREKA"
complement(5042..6183)
/standard_name="rrn T1r2 terminator"
complement(5042..6422)
/organism="Escherichia coli"
/db_xref="taxon:562"
complement(6183..6422)
/codon_start=1
/transl_table=11
/product="beta-galactosidase alpha peptide"
/protein_id="AAF78196.1"
/db_xref="GI:8650404"
/transl_table="WTMTTPSLHACRSTLEDPRVSSNSLAVIQRDMENPGVTQLN
RLAHPPEPASWNSSEARTDPSQDLRLNGECREL"

BASE COUNT 1926 a 1478 c 1345 g 1984 t
ORIGIN
Query Match      22.5%; Score 404; DB 12; Length 6733;
Best Local Similarity 96.5%; Pred. No. 5e-84;
Matches 413; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1267 aagatcgacctgcagcgatgcgaagcttggtcccaactgaacccatgccgaactcagaag 1326
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 5984 AGAATTTGCTGGCGGCACTAGCGCGGTGTCACCACTGACCCCATGCCAATCTCAGAAG 5925

QY 1327 tgaacgcgcgtagcgccgcatggtgtagtgtggtgtcccatcgagaagtaggaaactgcc 1386
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 5924 TGAACGCCGCTAGCGCGCATGTAGTGTGGGTCTCCCATCGAGAGTAGGGAATGCC 5865

QY 1387 aggcatacaataaagaaagcctcagtcgaaagactgggaccttcgtttatctgtgtc 1446
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 5864 AGGCATCAATTAAGAAAGAAAGCTCAGTGAAGAAAGCTGGGCTTTGCTTTATCTGTGT 5805

```

QY 1447 ttgtcgtgaaagcgtctcctgagtaggacaatccgcggagcgagatttgaaagtgcg 1506
|||||
Db 5804 TTGTCGCTGAAGCGCTCTCTGAGTAGAGACAATCCGCCGGAGCGGATTGAAGCTTGCG 5745
QY 1507 aagcaagcgccggagaggtggcgagcgccgcataaactgcaagcatcaatt 1566
|||||
Db 5744 AAGCAAGCGCCGGAGGCTGGCGGAGAGCGCCGCCATTAACCTGCAGCATCAAAATP 5685
QY 1567 aagcagaagcgcatccctgaagcgtcttcttgctttctacaactctctcgtcgc 1626
|||||
Db 5684 AAGCAGAAAGCCATCTCTGAGAGATGGCTTTTGCGCTTCTACAAACTCTCTGCTGTC 5625
QY 1627 atatctcaagcattcccccacagaatacgttaactagcctcgttttgcatacagaagaa 1686
|||||
Db 5624 ATATCTACAAGCCATCCGCCACAGATACGTAACCTAGCCTCTTTTGATCAGAGAAA 5565
QY 1687 gcagctat 1694
|||||
Db 5564 GCAGCTGT 5557

RESULT 13

LOCUS AF216803 6984 bp DNA circular SYN 01-DEC-2000
DEFINITION Shuttle vector pDL276, complete sequence.
ACCESSION AF216803
VERSION AF216803.1 GI:10179615
KEYWORDS
SOURCE Shuttle vector pDL276.
ORGANISM artificial sequence; vectors.

REFERENCE 1 (bases 1 to 6984)
AUTHORS Dunny,G.M., Lee,I.N. and Leblanc,D.J.
TITLE Improved electroporation and cloning vector system for
gram-positive bacteria
JOURNAL Appl. Environ. Microbiol. 57 (4), 1194-1201 (1991)
MEDLINE 91282471
PUBMED 1903518
REFERENCE 2 (bases 1 to 6984)
AUTHORS Dunny,G.M. and Bae,T.
TITLE Direct Submission
JOURNAL Submitted (15-DEC-1999) Microbiology, University of Minnesota, Box
196 UMC, 1460 Mayo Building, 420 Delaware St. SE, Minneapolis, MN
55435, USA

FEATURES

source
1. .6984
/organism="Shuttle vector pDL276"
/db_xref="taxon:11474"
/note="Gram positive-Gram negative shuttle vector"
1746..2540
/codon_start=1
/transl_table=11
/product="aminoglycoside transferase"
/protein_id="AAG13816.1"
/db_xref="GI:10179616"

CDS

/translation="MAKKRISPELKKLIEKYRCVKDTESGSPAKYKLVGENENTLYLK
MTDSRYKGTIVDEREKDMLLEGLPVKYLHERHDSNLSLMDADGVLCSEY
EDEOSPEKIIELYAEICIRLFHSIDSDPYTNSLDRLAEYLILNNDLADYCEKME
EDTPEDPRELDELKTEKPEELVSHGDLGDSNIFVKGVSFIDIGRSGRADKMW
YDIACVRSIRREDIGEQYVELFPDLGKIKPMEKIKYIILDELFL"
BASE COUNT 1869 a 1593 c 1510 g 2012 t
ORIGIN

Query Match 22.5%: Score 404: DB 12: Length 6984:

Best Local Similarity 96.5%: Pred. No. 5e-84: Mismatches 15: Indels 0: Gaps 0:

Matches 413: Conservative 0: Mismatches 15: Indels 0: Gaps 0:
QY 1267 agagtcgacactgacgcatgcaagcttggtccacactgacccatgacgcaactcagaag 1326
|||||
Db 6235 AGAATTTGCTGGCGGAGTAGGCGGTGTGCCACTGTACCCCATGCCCACTCAGAA 6176

QY 1327 tgaacgcgtaagcgcgcatgtagtgggtctcccatgacgaagtaggaactgcc 1386
|||||
Db 6175 TGAACGCCCGTAGCGCGCATGTGTGGGTCTCCCATGCAAGTAGGGAATCGCC 6116
QY 1387 aggcataaataaagaaagcgtcaatgcgaagacggttcgtttatctgtgtc 1446
|||||
Db 6115 AGGCATCAAAATAAACAAGGCTCACTGCAAAAGCTGGCTCTTTTATCTGTTGT 6056
QY 1447 ttgtcgtgaaagcgtctcctgagtaggacaatccgcggagcgagatttgaaagtgcg 1506
|||||
Db 6055 TTGTCGCTGAAGCGCTCTCTGAGTAGAGACAATCCGCCGGAGAGGAGATTGAAGCTTGCG 5996
QY 1507 aagcaagcgccggagaggtggcgagcgccgcataaactgcaagcatcaatt 1566
|||||
Db 5995 AAGCAAGCGCCGGAGGCTGGCGGAGAGCGCCGCCATTAACCTGCAGCATCAAAATP 5936
QY 1567 aagcagaagcgcatccctgaagcgtcttcttgctttctacaactctctcgtcgc 1626
|||||
Db 5935 AAGCAGAAAGCCATCTCTGAGAGATGGCTTTTGCGCTTCTACAAACTCTCTGCTGTC 5876
QY 1627 atatctcaagcattcccccacagaatacgttaactagcctcgttttgcatacagaagaa 1686
|||||
Db 5875 ATATCTACAAGCCATCCGCCACAGATACGTAACCTAGCCTCTTTTGATCAGAGAAA 5816
QY 1687 gcagctat 1694
|||||
Db 5815 GCAGCTGT 5808

RESULT 14

LOCUS SYNPTL61T 8565 bp DNA circular SYN 24-MAY-1994
DEFINITION Cloning vector pTL61T, complete sequence.
ACCESSION M29896
VERSION M29896.1 GI:209190
KEYWORDS
SOURCE Cloning vector; synthetic sequence.
ORGANISM Synthetic plasmid pTL61T (natural host E. coli).
artificial sequence; vectors.

REFERENCE 1 (bases 1 to 8565)
AUTHORS Lim,T. and Plier,R.
TITLE Improved vector system for constructing transcriptional fusions
that ensures independent translation of lacZ
JOURNAL J. Bacteriol. 172, 1077-1084 (1990)
MEDLINE 90130263
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
by T.G.Lim, 15-NOV-1989.
Authors indicate an uncharacterized mutation between bp 3274-3379,
which alters a EcoRI site normally present in the E. coli lacZ gene.

FEATURES

source
1. .8565
/organism="unidentified cloning vector"
/sub_species="Cloning vector pTL61T"
/db_xref="taxon:45196"
1..92
/note="multiple cloning site"
116..192
/note="RnaaseIII processing site"
238..239
/organism="Cloning vector pTL61T"
256..3330
/note="beta-d-galactosidase"
/codon_start=1
/transl_table=11
/protein_id="AA72803.1"
/db_xref="GI:209191"

/translation="MTMTDSLAAYVLRDMDENPGYTOINRLAAMPSPASRNSEAR
TDRPSOQLSLNGEMRFAPFAPEAVPESWLECDLPEADFTVVPSPNQWHGIDAPIT
NVTYPTVNPVPEPTENPGCYSLTFNNVDESLGOFRIIFDGVNSAFHLMCNLRW
GYGDSRLPSEFDLSAFLRAGENRLAAVYLIRSDSGSYLEDODMMSGIFRDVSLHK
PTTQISDFHATRFNDRAVLEAEVOMCGELRDYLRVTVALMVGEOVAAGTAAPG
GELIDRGYADRVTLRLNVEPKLMSAETPLTAYAVELTHADCTLIEAFCQDVCPR
EVRIENGLILLNGKPLLRGVNRHEHPLHGVOMDEQTNVODILMKQNFNAVPCSH

```

LOCUS       CYPGCG1408                9113 bp    DNA    circular SYN 01-DEC-2000
DEFINITION  Promoter-probe vector pCG1408, complete sequence.
ACCESSION   U21228
VERSION     U21228.1  GI:885956
KEYWORDS
SOURCE      .
  ORGANISM  .
            Cloning vector pCG1408.
            Cloning vector pCG1408
            artificial sequence; vectors.
            1 (bases 1 to 9113)
REFERENCE   David,M., Bartz,S.L., Weaver,J.D. and Kelley,J.L.
            Isolation and Characterization of Promoter Sequences from
            Clavibacter xyli subsp. cynodontis
            Unpublished
            2 (sites)
AUTHORS     Taylor,J., Stearman,R.S. and Urrutani,B.B.
            Development of a native plasmid as a cloning vector in Clavibacter
            xyli subsp. cynodontis
            Plasmid 29 (3), 241-244 (1993)
JOURNAL     93361581
MEDLINE
REFERENCE   3 (bases 1 to 9113)
AUTHORS     David,M.
            Direct Submission
            Submitted (17-FEB-1995) Michael David, Crop Genetics, 10150 Old
            Columbia Rd, Columbia, MD 21046, USA, 21046
FEATURES
  source
    1..9113

```

```

terminator
1185..1228
/note="r1"
terminator
1360..1388
/note="r2"
terminator
1684..1726
/note="r1"
terminator
1858..1885
/note="r2"
2702..3562
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AA070411.1"
/db_xref="GI:885958"
/transl_table="MSI0HFRVALLPFFAFCPLVFAHPETLVKVKDAEDQLGARVGI
IELDNSKRIIESEPRERFPMWSFTFKVLLCGAVLSRVDAQEDQLGRGIRHSQNDLVE
YSPVEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGKELTAFLAHNGSDVTRLE
DSWEELIKDITPRDRTTPMAAMATITRLKLTIGELTGLASRQOLIMMEADKVAQPR
LRSAIPACMFILADKSGAGSRGIIALGDPGRKSRIVITTTGSAQTMERRRQQLA
ELGASLIKHW"
4084..4508
/note="E. coli origin of replication"
5465..8413
/note="Clavibacter origin of replication"
/citation[1]
/citation[2]
source
join(8414..9113,1..5464)
/organism="Cloning vector pKK232-8"
/note="region derived from vector pKK232-8, GenBank
Accession Number U13859"
/db_xref="taxon:74582"
8947..8990
/note="r1"
terminator

```


